Journal of Computational and Applied Mathematics 427 (2023) 115158

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

A high-order numerical scheme for stochastic optimal control problem ☆

Yunzhang Li

Research Institute of Intelligent Complex Systems, Fudan University, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 28 August 2022
Received in revised form 30 January 2023

MSC:
65M60
65M15
65Z05

Keywords:
Stochastic optimal control
Local discontinuous Galerkin method
Fully nonlinear second-order PDEs
Optimal error estimates

ABSTRACT

In this paper, we propose a local discontinuous Galerkin (LDG) method for fully nonlinear second-order PDEs in multiple space dimensions, which can serve as a high-order numerical scheme for stochastic optimal control problems. The optimal error estimates are obtained for smooth solutions. Some numerical examples are given to display the performance of the LDG method.

1. Introduction

In this paper we consider the following fully nonlinear second-order PDEs in multiple space dimensions:

$$\begin{cases} u_t = F(D^2u, \nabla u, u, x, t), & (x,t) \in D \times (0,T], \\ u(x,0) = u_0(x), & x \in D. \end{cases} \tag{1.1}$$

We shall consider the periodic boundary condition: $x \in D \triangleq \mathbb{T}^d$, where $\mathbb{T}^d$ is the d-dimensional torus ($d \geq 1$). Notice that the assumption of periodic boundary condition is for simplicity of exposition only and is not essential: the method can be easily designed for arbitrary domain and for non-periodic boundary condition. Fully nonlinear second-order PDEs (1.1) arise from many scientific and engineering fields such as astrophysics, antenna design, geostrophic fluid dynamics, image processing, materials science, mathematical finance, mesh generation, meteorology, optimal transport, and stochastic control (cf. [1–4] and the references therein); they are a class of PDEs which are very difficult to analyse and even more challenging to approximate numerically. The goal of this paper is to develop a local discontinuous Galerkin (LDG) method for (1.1) with proving its optimal error estimates and applying the scheme to stochastic control problems.

The first DG method was introduced by Reed and Hill to solve a steady linear transport problem [5] in 1973. The LDG method was introduced by Cockburn and Shu [6] for solving quasi-linear convection diffusion equations, which was motivated by the successful numerical experiments of Bassi and Rebay [7] for the compressible Navier–Stokes equations. This scheme is an extension of the discontinuous Galerkin (DG) method developed by Cockburn et al. [8–11] for nonlinear

☆ This work was supported by the Chenguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission, and China National Postdoctoral Program for Innovative Talents (Grant No. BX20200096), and China Postdoctoral Science Foundation (Grant No. 2021M690703).

E-mail address: li_yunzhang@fudan.edu.cn.

https://doi.org/10.1016/j.cam.2023.115158
0377-0427/

hyperbolic systems and shares with the DG method its advantage and flexibility: (1) it allows for easy design of high-order approximations, which in turn allows for efficient p-adaptivity; (2) it is flexible on complicated geometries, which allows for efficient h-adaptivity; (3) it is local in data communications, which allows for efficient parallel implementations. The LDG method also inherits many good properties such as conservation, L2-boundedness, optimal error estimates and super-convergence results from the conventional DG scheme, which makes it quite attractive.

Due to their fully nonlinear structures disenable us to perform a direct integration by parts, fully nonlinear PDEs do not have variational formulations in general. So in fact, there was no weak solution concept for fully nonlinear PDEs until Crandall and Lions [12] introduced the notion of viscosity solutions for fully nonlinear first-order PDEs. Then their notion and theory were quickly extended to fully nonlinear second-order PDEs. The non-variational structure prevents the applicability of standard Galerkin type methods such as DG methods. On the other hand, we have to say that, to approximate the low-regularity viscosity solutions, it is natural to use totally discontinuous piece-wise polynomial functions (i.e., DG functions) due to their flexibility and the larger approximation spaces.

Concerning the study of DG schemes for the fully nonlinear first-order time-dependent Hamilton–Jacobi (HJ) equations, let us first mention that Hu and Shu [13] developed a DG scheme based on the classical Runge–Kutta DG method for solving conservation laws satisfied by the first-order derivative of the solution. Later, Li and Shu [14] reinterpreted the method in [13] by using a curl-free subspace for the DG method in the two-dimensional case, resulting in a significant simplification in implementation with a reduced cost. After that, Cheng and Shu [15] proposed a DG method for directly solving HJ equations without going through the derivative of the solution. Also, Yan and Osher [16] designed a direct LDG method for approximating HJ equations. More recently, Xiong, Shu and Zhang [17] obtained a priori optimal error estimates for the DG and LDG methods in [15,16] for smooth solutions.

In regard to the study of the DG type schemes for fully nonlinear second-order PDEs, Feng and Lewis [18] first designed a class of nonstandard mixed interior penalty DG method, which works well provided that the viscosity solutions belong to $C^0\left(\overline{D}\right) \cap H^1(D)$ and the polynomial degree k is greater than or equal to 1. After that, they [19,20] presented non-standard LDG methods by using multiple approximations of first and second-order derivatives for fully nonlinear second-order PDEs in both one and multi-dimensional cases. But none of these articles gives the theoretical error estimates for numerical solutions. It is also worth noting that Brenner and Neilan et al. investigated some finite element methods for approximating smooth solutions of fully nonlinear PDEs with convergence analysis (see e.g. [21,22]).

In this paper, inspired by [16,19,20] we present an LDG method for directly solving fully nonlinear second-order PDEs (1.1) in multiple space dimensions by using the alternating numerical flux. It should be pointed out that our LDG method can be regarded as a special case of the non-standard LDG methods proposed in [19,20], in which the error estimates were not analysed. We drop the numerical moment term to make our scheme more easier to implement compared the LDG methods in [19,20]. To fill the gap in numerical analysis of the DG type methods for fully nonlinear second-order PDEs, we also follow and generalize the techniques in [17,23–25] to prove a priori L^2-error estimates for classical strong solutions with enough smoothness and integrability. To theoretically analyse our numerical scheme, we need to make the following uniformly parabolic assumption.

($\mathbf{P}_\delta$) The function $F:\mathbb{R}^{d\times d}\times\mathbb{R}^d\times\mathbb{R}\times D\times[0,T]\longrightarrow\mathbb{R}$ is in C_b^2, and there exists a constant $\delta>0$ such that for each $(x,t)\in D\times[0,T]$, it holds

$$F_{\mathbb{P}}(\Theta) := \left[\frac{\partial F}{\partial p^{ml}}(\Theta)\right]_{d\times d} \geq \delta I_d,$$

where $\Theta := (\mathbb{P},\vec{v},u,\cdot)(x,t) \in \mathbb{R}^{d\times d}\times\mathbb{R}^d\times\mathbb{R}\times D\times[0,T]$, with u being the unique exact solution of (1.1), $\vec{v}=[v^m]_{d\times 1} := \nabla u = \left[\frac{\partial u}{\partial x^m}\right]_{d\times 1}$ and $\mathbb{P}=\left[p^{ml}\right]_{d\times d} := D^2u = \left[\frac{\partial v^l}{\partial x^m}\right]_{d\times d}$.

Note that the non-degenerate and smooth condition ($\mathbf{P}_\delta$) is used only in the theoretical proof of error estimates, not in the implementation of our LDG method. In Sections 4 and 5, some of the numerical experiments do not fall in the scope of the present set-up. However the numerical results show that the algorithm is robust enough and works well even in the degenerate and non-smooth cases. The paper is organized as follows. In Section 2, we introduce notations and definitions used later in this paper. In Section 3, we state the LDG method for fully nonlinear second-order PDEs and present a priori error estimates for smooth solutions. In Section 4, we present a series of numerical results to validate the LDG method. In Section 5, we use our numerical method to solve stochastic optimal control problems. In Section 6, we give the technical proofs of the convergence results for the LDG method. Finally, in Section 7, concluding remarks are given.

2. Notations and definitions

2.1. One-dimensional case

We denote the mesh by $I_j=\left[x_{j-\frac{1}{2}},x_{j+\frac{1}{2}}\right]$, for $j=1,\ldots,N$. The nodes are denoted by $\left\{x_{j+\frac{1}{2}},\ j=0,\ldots,N\right\}$ with $x_{\frac{1}{2}}=0$ and $x_{N+\frac{1}{2}}=1$. We define $x_j := \frac{1}{2}\left(x_{j-\frac{1}{2}}+x_{j+\frac{1}{2}}\right)$. The mesh size is $h_j := x_{j+\frac{1}{2}}-x_{j-\frac{1}{2}}$, with $h=\max_{1\leq j\leq N} h_j$

being the maximum mesh size. For a given function f, we denote $f\left(x^{\pm}_{j+\frac{1}{2}}\right)$ by $f^{\pm}_{j+\frac{1}{2}}$, where $x^{+}_{j+\frac{1}{2}} := \lim_{x\searrow x_{j+\frac{1}{2}}} x$ and $x^{-}_{j+\frac{1}{2}} := \lim_{x\nearrow x_{j+\frac{1}{2}}} x$. The finite element space is

$$V_h^k := \left\{ f \ :\ f \in P^k(I_j), \quad \text{for } x \in I_j,\ \ j = 1, \ldots, N \right\},$$

where $P^k(I_j)$ is the space of polynomials of total degree at most k. Note that functions in V_h^k might have discontinuities on an element interface.

2.2. Multi-dimensional case

We consider triangulations of $D = \mathbb{T}^d$ with $d \geq 2$, $\mathcal{T}_h = \{K\}$, made of nonoverlapping polyhedra completely covering D. The diameter of K is denoted by h_K and the maximum h_K, for $K \in \mathcal{T}_h$ is denoted by h. We require the triangulations $\mathcal{T}_h$ to be regular. The centroid of the triangular K is denoted by x_K. We denote the boundary of element K by ∂K. Let Γ_h denote the union of the boundary faces of elements $K \in \mathcal{T}_h$, i.e. $\Gamma_h := \bigcup_{K\in\mathcal{T}_h} \partial K$. The outward normal unit vector to ∂K is denoted by $\vec{n}_K$.

We use the notation

$$\left(f\,,\,\vec{g}\cdot\vec{n}_K\right)_{\partial K} := \int_{\partial K} f(x)\vec{g}(x)\cdot\vec{n}_K\,d\,\Gamma(x), \qquad \left(f\,,\,\vec{g}\cdot\vec{n}_K\right)_{\Gamma_h} := \sum_{K\in\mathcal{T}_h}(f\,,\,\vec{g}\cdot\vec{n}_K)_{\partial K},$$

for any functions $f, \vec{g}$ in the broken Sobolev space $H^{1,2}(\mathcal{T}_h)$, which is the space of functions that are elementwise in $H^{1,2}$ Sobolev space.

For a given function f, f^+ denotes the value of f evaluated from a predesignated "plus" side along an edge e, which is always the boundary of two neighbouring elements. For example, we could choose a fixed vector $\vec{\beta}$, which is not parallel with any element boundary, and then designate the "plus" side to be the side at the end of the arrow of the normal $\vec{n}_K$ with $\vec{n}_K\cdot\vec{\beta} > 0$. The definition of f^- is given in a similar way. For more details, we refer to [26].

The finite element spaces associated with the mesh $\mathcal{T}_h$ are of the form

$$V_h^k := \left\{ v\ :\ v|_K \in \mathbf{V}^k(K),\ \ \text{for } K \in \mathcal{T}_h \right\},$$

where $\mathbf{V}^k(K)$ denotes the local space on the element K. For triangular meshes, we let $\mathbf{V}^k(K) = P^k(K)$, where $P^k(K)$ is the space of polynomials of total degree at most k. For Cartesian meshes, we let $\mathbf{V}^k(K) = Q^k(K)$, where $Q^k(K)$ is the space of tensor product of polynomials of degrees at most k in each variable.

For the function $F(\mathbb{P}, \vec{v}, u, x, t)$, we denote

$$F_{\mathbb{P}\mathbb{P}} := \left[\frac{\partial^2 F}{\partial p^{ml}\partial p^{ij}}\right]_{d\times d\times d\times d}, \qquad F_{\mathbb{P}} := \left[\frac{\partial F}{\partial p^{ml}}\right]_{d\times d}, \qquad F_{\vec{v}} := \left[\frac{\partial F}{\partial v^m}\right]_{d\times 1}, \qquad F_u := \frac{\partial F}{\partial u},$$

where $(\mathbb{P}, \vec{v}, u, x, t) \in \mathbb{R}^{d\times d} \times \mathbb{R}^d \times \mathbb{R} \times D \times [0, T]$. The norm of a $d_1 \times d_2$ matrix y is given by $|y| := \sqrt{trace(yy^{\mathrm{T}})}$. We write $\|\cdot\|$ and $\|\cdot\|_{H^m}$ for the $L^2(D)$-norm and Sobolev norm $\|\cdot\|_{H^{m,2}(D)}$, respectively. We denote by C_b^m the space of continuous functions with bounded m-order derivatives.

Throughout the paper, by saying that a vector-valued or matrix-valued function belongs to a function space, we mean all the components belong to that space. By $C > 0$, we denote a generic constant, which in particular does not depend on the discretization width h and possibly changes from line to line. When there is no ambiguity, we omit the argument t in the proofs for simplicity of notations.

3. The LDG method and error estimates

3.1. One-dimensional case

To illustrate the idea of our scheme, for the sake of easy presentation, we first consider the model problem with $d = 1$:

$$\begin{cases} u_t = F(u_{xx}, u_x, u, x, t), & (x, t) \in \mathbb{T} \times (0, T], \\ u(x, 0) = u_0(x), & x \in \mathbb{T}. \end{cases} \tag{3.1}$$

3.1.1. The LDG method in one space dimension

As a special class of the DG methods, the main technique of the LDG method is to rewrite (3.1) into an equivalent system containing only first-order spatial derivatives, which is further discretized by the standard DG method with correct definition of numerical fluxes. To do this, firstly, we rewrite the problem as a first-order system:

$$\begin{cases} u_t = F(p, v, u, x, t), & \text{(a)} \\ v = u_x, & \text{(b)} \\ p = v_x. & \text{(c)} \end{cases} \tag{3.2}$$

The LDG method for (3.1) is now obtained by simply discretizing the above system with the DG method. We seek an approximation (u_h, v_h, p_h) to the exact solution (u, v, p) such that for any $t \in [0, T]$, $(u_h, v_h, p_h)(\cdot, t)$ belongs to the finite dimensional space V_h^k. In order to determine the approximate solution (u_h, v_h, p_h), we first note that by multiplying (3.2a), (3.2b), (3.2c) with arbitrary smooth functions z_u, z_v, z_p, respectively, and integrating over I_j with $j = 1, 2, \ldots, N$, we get, after a simple formal integration by parts in(3.2b) and (3.2c),

$$\begin{cases} \int_{I_j} u_t\,(x,t)\, z_u(x)dx = \int_{I_j} F(p, v, u, \cdot)(x,t) z_u(x)dx, \\ \int_{I_j} v\,(x,t)\, z_v(x)dx = -\int_{I_j} u\,(x,t)\,(z_v)_x\,(x)\,dx + \left[u\,(\cdot,t)\, z_v\right]\Big|_{x^+_{j-\frac12}}^{x^-_{j+\frac12}}, \\ \int_{I_j} p\,(x,t)\, z_p(x)dx = -\int_{I_j} v\,(x,t)\,(z_p)_x\,(x)\,dx + \left[v\,(\cdot,t)\, z_p\right]\Big|_{x^+_{j-\frac12}}^{x^-_{j+\frac12}}. \end{cases}$$

Next, in the above weak formulation, we replace the smooth function (z_u, z_v, z_p) with the test function $(z_{h,u}, z_{h,v}, z_{h,p})$ in the finite element space V_h^k and the exact solution (u, v, p) with the approximation (u_h, v_h, p_h). Since the functions in V_h^k might have discontinuities on an element interface, we must also replace the boundary terms $u(x_{j+\frac12}, t)$ and $v(x_{j+\frac12}, t)$ with the numerical fluxes $\widehat{u}_{j+\frac12}(t)$ and $\widehat{v}_{j+\frac12}(t)$, respectively, which are defined by

$$\widehat{u}_{j+\frac12}\,(t) := u_h\left(x^-_{j+\frac12}, t\right), \qquad \widehat{v}_{j+\frac12}\,(t) := v_h\left(x^+_{j+\frac12}, t\right), \tag{3.3}$$

for $j = 0, 1, \ldots, N$. Note that, by periodicity, we have

$$\widehat{u}_{\frac12} = \widehat{u}_{N+\frac12} \quad \text{and} \quad \widehat{v}_{N+\frac12} = \widehat{v}_{\frac12}.$$

Thus, the approximate solution given by the LDG method is defined as the solution of the following weak formulation:

$$\begin{cases} \int_{I_j} (u_h)_t\,(x,t)\, z_{h,u}(x)dx = \int_{I_j} F(p_h, v_h, u_h, \cdot)\,(x,t)\, z_{h,u}(x)dx, & \text{(a)} \\ \int_{I_j} v_h\,(x,t)\, z_{h,v}(x)dx = -\int_{I_j} u_h\,(x,t)\,(z_{h,v})_x(x)dx + u_h\left(x^-_{j+\frac12}, t\right) z_{h,v}\left(x^-_{j+\frac12}\right) - u_h\left(x^-_{j-\frac12}, t\right) z_{h,v}\left(x^+_{j-\frac12}\right), & \text{(b)} \\ \int_{I_j} p_h\,(x,t)\, z_{h,p}(x)dx = -\int_{I_j} v_h\,(x,t)\,(z_{h,p})_x(x)dx + v_h\left(x^+_{j+\frac12}, t\right) z_{h,p}\left(x^-_{j+\frac12}\right) - v_h\left(x^+_{j-\frac12}, t\right) z_{h,p}\left(x^+_{j-\frac12}\right), & \text{(c)} \end{cases} \tag{3.4}$$

for any $(z_{h,u}, z_{h,v}, z_{h,p})$ in V_h^k. We define the initial value of the LDG method to be the L^2-projection of u^0, i.e., $u_h(\cdot, 0) := \mathcal{P}u^0$.

Remark 3.1. The choice of $(\widehat{u}, \widehat{v})$ in (3.3) is called alternating flux, which is essential for the proof of optimal error estimates. We can also define the numerical flux in an alternating way as

$$\widehat{u}_{j+\frac12}(t) := u_h(x^+_{j+\frac12}, t) \quad \text{and} \quad \widehat{v}_{j+\frac12}(t) := v_h(x^-_{j+\frac12}, t).$$

For simplicity of notation, for $j = 1, 2, \ldots, N$ and piece-wisely smooth functions f and g, we define

$$H_j^{\pm}(f, g) := -\int_{I_j} f\,(x)\, g_x(x)\,dx + f\left(x^{\pm}_{j+\frac12}\right) g\left(x^-_{j+\frac12}\right) - f\left(x^{\pm}_{j-\frac12}\right) g\left(x^+_{j-\frac12}\right). \tag{3.5}$$

Then the LDG method (3.4) reads, for any $(z_{h,u}, z_{h,v}, z_{h,p})$ in V_h^k,

$$\begin{cases} \int_{I_j} (u_h)_t\,(x,t)\, z_{h,u}(x)dx = \int_{I_j} F(p_h, v_h, u_h, \cdot)\,(x,t)\, z_{h,u}(x)dx, & \text{(a)} \\ \int_{I_j} v_h\,(x,t)\, z_{h,v}(x)dx = H_j^-\left(u_h(\cdot,t),\, z_{h,v}\right), & \text{(b)} \\ \int_{I_j} p_h\,(x,t)\, z_{h,p}(x)dx = H_j^+\left(v_h(\cdot,t),\, z_{h,p}\right), & \text{(c)} \end{cases} \tag{3.6}$$

where $j = 1, 2, \ldots, N$ and $u_h(\cdot, 0) := \mathcal{P}u^0$. As a semi-discrete scheme, the algorithm (3.6) is not difficult for numerical implementation. In fact, given u_h, one first uses (3.6b) to locally solve for v_h, then uses (3.6c) to locally solve p_h, and finally uses (3.6a) to locally solve for the update of u_h.

3.1.2. Optimal error estimates in one space dimension

Let u be the exact solution of the problem (3.1) with $(v, p) := (u_x, u_{xx})$. The numerical solution (u_h, v_h, p_h) is calculated by the semi-discrete LDG scheme (3.6). We have the following convergence result for our numerical methods, and its proof is given in Section 6.2.

Theorem 3.1. *Let Assumption* $(\mathbf{P}_\delta)$ *hold, and the considered fully nonlinear PDEs (3.1) admit a unique solution* $u \in L^\infty((0,T), H^{k+2}(D))$. *If the finite element space* V_h^k *is the piece-wise polynomial space of degree* $k \geq 2$, *then for small enough* h, *there holds the following optimal error estimate*

$$\sup_{0\leq t\leq T} \|u(\cdot,t) - u_h(\cdot,t)\| + \left(\int_0^T \|u_x(\cdot,t) - v_h(\cdot,t)\|^2 \, dt\right)^{\frac{1}{2}} \leq Ch^{k+1},$$

where the positive constant C *depends on* T, δ, k, $\|u\|_{L^\infty((0,T),H^{k+2}(D))}$ *and the bounds of* F_p, F_v, F_u *and* F_{pp}.

Remark 3.2. Note that the assumption $k \geq 2$ in Theorem 3.1 is only for the theoretical analysis of optimal error estimates, and is not necessary in the numerical implementation. The numerical experiments in Section 4 indicate that our scheme still works well when $k < 2$. The similar conclusion also holds in the multi-dimensional case.

3.2. Multi-dimensional case

In this subsection, we generalize the scheme discussed in the previous subsection to multiple space dimensions $x = (x^1, \ldots, x^d) \in D = \mathbb{T}^d$.

3.2.1. The LDG method in multiple space dimension

Similar to the one-dimensional case, to define the LDG method in multiple space dimensions, we first rewrite (1.1) as a first order system:

$$\begin{cases} u_t = F(\mathbb{P}, \vec{v}, u, x, t), \\ v^i = \dfrac{\partial u}{\partial x^i}, \qquad i = 1, \ldots, d, \\ p^{ml} = \dfrac{\partial v^l}{\partial x^m}, \qquad m, l = 1, \ldots, d. \end{cases}$$

We denote $\vec{v} = \left[v^i\right]_{d\times 1}$ and $\mathbb{P} = \left[p^{ml}\right]_{d\times d}$. We proceed exactly as in the one-dimensional case. This time, however, the integrals are made on each element K of the triangulation $\mathcal{T}_h$. For arbitrary smooth functions z_u, $\vec{z}_v = (z_v^1, \ldots, z_v^d)^{\mathrm{T}}$ and z_p, and for all $d \times d$ matrices $A_K = \left[A_K^{ml}\right]_{d\times d}$, it holds that

$$\begin{aligned} &\int_K u_t(x,t) z_u(x) dx = \int_K F(\mathbb{P}, \vec{v}, u, \cdot)(x,t) z_u(x) dx, \\ &\int_K \vec{v}(x,t) \cdot \vec{z}_v(x) dx = -\int_K u(t,x) \nabla \cdot \vec{z}_v(x) dx + \left(u|^{intK}(t,\cdot), \vec{z}_v|^{intK} \cdot \vec{n}_K\right)_{\partial K}, \\ &\int_K \sum_{m,l=1}^d A_K^{ml} p^{ml}(x,t) z_p(x) dx = -\int_K A_K \vec{v}(x,t) \cdot \nabla z_p(x) dx + \left(A_K \vec{v}|^{intK} \cdot \vec{n}_K, z_p|^{intK}\right)_{\partial K}, \end{aligned} \tag{3.7}$$

where $f|^{intK}$ denotes the value of the considered function f evaluated from inside the element K.

Next, we replace the smooth functions z_u, $\vec{z}_v$ and z_p by test functions $z_{h,u}$, $\vec{z}_{h,v} = (z_{h,v}^1, \ldots, z_{h,v}^d)^{\mathrm{T}}$ and $z_{h,p}$, respectively, in the finite element space V_h^k, and the exact solution $(u, \vec{v}, \mathbb{P})$ by the approximate solution $(u_h, \vec{v}_h, \mathbb{P}_h)$. Here we denote $\vec{v}_h := \left[v_h^i\right]_{d\times 1}$ and $\mathbb{P}_h := \left[p_h^{ml}\right]_{d\times d}$. We again need to carefully choose numerical fluxes for the boundary terms resulting from the procedure of integration by parts. We still choose the alternating numerical flux $\widehat{u} = u_h^-$ and $\widehat{\vec{v}} = \vec{v}_h^+$, where the "plus" side and "minus" side are defined in Section 2.2. Then we obtain the LDG scheme for fully nonlinear second-order PDEs in multi-dimensional case, which is the following weak formulation on each K of the triangulation $\mathcal{T}_h$: for each fixed $t \in [0,T]$, it holds that for all $d \times d$ matrices $A_K = \left[A_K^{ml}\right]_{d\times d}$

$$\begin{cases} \displaystyle\int_K (u_h)_t (x,t) z_{h,u}(x) dx = \int_K F(\mathbb{P}_h, \vec{v}_h, u_h, \cdot)(x,t) z_{h,u}(x) dx, & \text{(a)} \\ \displaystyle\int_K \vec{v}_h(x,t) \cdot \vec{z}_{h,v}(x) dx = H_K^-\left(u_h(\cdot,t), \vec{z}_{h,v}\right), & \text{(b)} \\ \displaystyle\int_K \sum_{m,l=1}^d A_K^{ml} p_h^{ml}(x,t) z_{h,p}(x) dx = H_K^+\left(A_K \vec{v}_h(\cdot,t), z_{h,p}\right), & \text{(c)} \end{cases} \tag{3.8}$$

with $u_h(\cdot,0) := \Pi u^0$, where Π is the standard L^2-projection onto V_h^k, and the bilinear functionals are defined as

$$H_K^+(\vec{g},f) := -\int_K \vec{g}(x)\cdot\nabla f(x)dx + \left(\vec{g}^{\,+}\cdot\vec{n}_K, f|^{intK}\right)_{\partial K},$$

$$H_K^-(f,\vec{g}) := -\int_K f(x)\,\nabla\cdot\vec{g}(x)dx + \left(f^-, \vec{g}|^{intK}\cdot\vec{n}_K\right)_{\partial K}.$$

Remark 3.3. Note that the scheme (3.8c) stands for all $d\times d$ matrices $A_K = \left[A_K^{ml}\right]_{d\times d}$, which is designed based on (3.7) resulted form the Green formula. In practice, one can compute p_h^{ml} by setting $A_K^{ij} := \delta_{ml}$ for each fixed $l, m \in \{1,\dots,d\}$, where δ_{ml} equals to 1 if $(i,j) = (m,l)$, and equals to 0 otherwise.

3.2.2. Error estimates in multiple space dimensions

Now we give a priori error estimates for the approximation $(u_h, \vec{v}_h, \mathbb{P}_h)$ given by the LDG method (3.8) in multi-dimensional case. Let u be the exact solution of the problem (1.1) with $(\vec{v},\mathbb{P}) := (\nabla u, D^2u)$. Without loss of generality, we study the error estimates in two dimensions ($d = 2$). The proof of the following theorem is presented in Section 6.3.

Theorem 3.2. *Let Assumption ($\mathbf{P}_\delta$) hold and the considered fully nonlinear PDEs (1.1) admit a unique solution $u \in L^\infty((0,T),H^{k+2}(D))$. For the two-dimensional Cartesian meshes with Q^k elements and $k\ge 3$, there holds the following optimal error estimates*

$$\sup_{0\le t\le T}\|u(\cdot,t)-u_h(\cdot,t)\| + \left(\int_0^T \|\nabla u(\cdot,t)-\vec{v}_h(\cdot,t)\|^2\,dt\right)^{\frac{1}{2}} \le Ch^{k+1};$$

and for the two-dimensional triangular meshes with P^k elements and $k \ge 4$, there holds the following sub-optimal error estimates

$$\sup_{0\le t\le T}\|u(\cdot,t)-u_h(\cdot,t)\| + \left(\int_0^T \|\nabla u(\cdot,t)-\vec{v}_h(\cdot,t)\|^2\,dt\right)^{\frac{1}{2}} \le Ch^{k},$$

where the positive constant C depends on T, δ, k, $\|u\|_{L^\infty((0,T),H^{k+2}(D))}$ and the bounds of $F_{\mathbb{P}}$, $F_{\vec{v}}$, F_u and $F_{\mathbb{P}\mathbb{P}}$.

4. Numerical experiments

Up to now, we have taken the method of lines approach and have left time t continuous. According to the continuity and boundedness assumption of F, it can be shown that the method of line ODE

$$\varphi_t = L(\varphi,t),$$

admits a unique solution and we can use high-order Runge–Kutta methods to solve it. The third-order Runge–Kutta method that we use in this paper is given by

$$\varphi^{n+1} = \varphi^n + \frac{1}{9}\left(2\psi_1 + 3\psi_2 + 4\psi_3\right), \tag{4.1}$$

where

$$\begin{cases} \psi_1 &= \Delta t\, L\left(\varphi^n, t^n\right), \\ \psi_2 &= \Delta t\, L\left(\varphi^n + \frac{1}{2}\psi_1, t^n + \frac{1}{2}\Delta t\right), \\ \psi_3 &= \Delta t\, L\left(\varphi^n + \frac{3}{4}\psi_2, t^n + \frac{3}{4}\Delta t\right). \end{cases}$$

Next we are going to use the time discretization (4.1) to provide numerical experimental results for demonstrating the behaviour of our LDG scheme. Since the considered fully nonlinear PDEs (1.1) involve second-order spatial derivatives, in all experiments, we need to adjust the time step to $\Delta t \sim (\Delta x)^2$ to guarantee the stability for the explicit time discretization. Moreover, by setting $\Delta t \sim (\Delta x)^2$, the scheme in time is efficiently sixth-order with respect to h. In the future, we plan to investigate some implicit–explicit scheme to relax this small step restriction.

Although in the numerical analysis we assume that the coefficients of the fully nonlinear equations are bounded and smooth enough, it is worth trying to apply the LDG scheme to some nonlinear equations with unbounded and discontinuous coefficients in this section. And the numerical experiments shows that our LDG scheme works well in different cases.

4.1. One-dimensional case

In this subsection, one-dimensional problems are computed using our scheme.

Table 1
Accuracy on (4.2) with $T = 0.1$.

h	$k=0$		$k=1$		$k=2$	
	e_2	Order	e_2	Order	e_2	Order
$2\pi/10$	2.92E−01	–	3.97E−02	–	1.94E−03	–
$2\pi/20$	1.46E−01	1.00	9.94E−03	2.00	2.43E−04	3.00
$2\pi/40$	7.32E−02	1.00	2.48E−03	2.00	3.03E−05	3.00

4.1.1. Non-degenerate equation with smooth F in one space dimension

We first consider the non-degenerate fully-nonlinear second-order PDE

$$\begin{cases} u_t(x,t) &= \left[(u_{xx})^3 + u_{xx} + (u_x)^2 + u^2\right](x,t) + f(x,t), \\ u(x,0) &= \sin(x), \end{cases} \tag{4.2}$$

where we take

$$f(x,t) := \sin^3(x)\exp(-3t) - \exp(-2t),$$

such that the analytic solution of (4.2) is

$$u(x,t) = \sin(x)\exp(-t), \qquad (x,t) \in [0, 2\pi] \times [0, T]. \tag{4.3}$$

Here, we have

$$F(p, v, u, x, t) = p^3 + p + v^2 + u^2 + f(x,t),$$

which is a smooth function. Since

$$F_p(p, v, u, x, t) = 3p^2 + 1 \geq 1 > 0,$$

we know that (4.2) is a non-degenerate equation with $\delta = 1$.

The L^2-errors at the terminal time T

$$e_2 := \|u_h(\cdot, T) - u(\cdot, T)\|$$

and their convergence rates are listed in Table 1. We clearly observe $(k+1)$-th order of accuracy for P^k polynomials, which confirms the result in Theorem 3.1.

4.1.2. Degenerate equation with smooth F in one space dimension

Although we cannot give the error estimates for the degenerate problems, it is worth trying to apply the LDG scheme (3.6) to some degenerate equations. So the next example is a degenerate fully-nonlinear second-order PDE

$$\begin{cases} u_t(x,t) &= \left[(u_{xx})^3 + (u_x)^2 + u^2\right](x,t) + f(x,t), \\ u(x,0) &= \sin(x), \end{cases} \tag{4.4}$$

where we take

$$f(x,t) := \sin^3(x)\exp(-3t) - \sin(x)\exp(-t) - \exp(-2t),$$

such that the exact solution of (4.4) is (4.3).

Here, we have

$$F(p, v, u, x, t) = p^3 + v^2 + u^2 + f(x,t),$$

and

$$F_p(p, v, u, x, t) = 3p^2 \geq 0,$$

which indicates that (4.4) is a degenerate equation.

In this example, we take $T = 0.1$ and list the numerical results in Table 2, which shows that the optimal $(k+1)$-th order of accuracy still holds for the degenerate case.

4.1.3. Non-degenerate equation with non-smooth F in one space dimension

Next we consider the one-dimensional non-degenerate Hamilton–Jacobi–Bellman (HJB) equation:

$$\begin{cases} u_t(x,t) &= \min\left\{\frac{1}{2}u_{xx}, u_{xx}\right\}(x,t) + f(x,t), \\ u(x,0) &= \sin(x), \end{cases} \tag{4.5}$$

Table 2
Accuracy on (4.4) with $T = 0.1$.

h	$k = 0$		$k = 1$		$k = 2$	
	e_2	Order	e_2	Order	e_2	Order
$2\pi/10$	2.92E−01	–	4.08E−02	–	2.15E−03	–
$2\pi/20$	1.47E−01	0.99	1.03E−02	1.99	2.70E−04	2.99
$2\pi/40$	7.34E−02	1.00	2.58E−03	2.00	3.16E−05	3.10

Table 3
Accuracy on (4.5) with $T = 0.1$.

h	$k = 0$		$k = 1$		$k = 2$	
	e_2	Order	e_2	Order	e_2	Order
$2\pi/10$	2.89E−01	–	3.88E−02	–	1.93E−03	–
$2\pi/20$	1.45E−01	0.99	9.66E−03	2.00	2.43E−04	2.99
$2\pi/40$	7.27E−02	1.00	2.41E−03	2.00	3.03E−05	3.00

where we take

$$f(x,t) := -\sin(x)\exp(-t) + \max\left\{\frac{1}{2}\sin(x), \sin(x)\right\}\exp(-t),$$

such that the exact solution of (4.5) is (4.3).

Here, we have

$$F(p, v, u, x, t) = \min\left\{\frac{1}{2}p\,, p\right\} + f(x,t),$$

which is a non-smooth function. Note that

$$F_p(p, v, u, x, t) = \begin{cases} 1, & \text{if } p \le 0, \\ \dfrac{1}{2}, & \text{if } p > 0. \end{cases}$$

Since $F_p \ge \frac{1}{2} > 0$, we know that (4.5) is a non-degenerate equation with $\delta = \frac{1}{2}$.

Table 3 gives the errors of numerical solution at $T = 0.1$ with $0 \le k \le 2$. We clearly see that the LDG method has optimal $(k+1)$-th order of accuracy for the HJB equation with non-smooth F.

4.2. Multi-dimensional case

In this subsection, two-dimensional problems are approximated using our scheme with Cartesian meshes and triangular meshes.

4.2.1. Non-degenerate equation with smooth F in multiple space dimensions

We consider the two-dimensional non-degenerate fully nonlinear second-order PDE

$$\begin{cases} u_t(x,y,t) &=& \left[\dfrac{1}{2}(u_{xx})^3 + \dfrac{1}{2}\left(u_{yy}\right)^3 + \Delta u\right](x,y,t) + \sin^3(x+y)\exp(-6t), \\ u(x,y,0) &=& \sin(x+y). \end{cases} \tag{4.6}$$

Here, we have

$$F(\mathbb{P}, \vec{v}, u, x, y, t) = \left[\frac{1}{2}(p_{11})^3 + \frac{1}{2}(p_{22})^3 + p_{11} + p_{22}\right] + \sin^3(x+y)\exp(-6t),$$

which is a smooth function. Since

$$F_{\mathbb{P}}(\mathbb{P}, \vec{v}, u, x, y, t) = \begin{bmatrix} \dfrac{3}{2}p_{11}^2 + 1 & 0 \\ 0 & \dfrac{3}{2}p_{22}^2 + 1 \end{bmatrix} \ge I_{2\times 2},$$

we know that (4.6) is a non-degenerate equation with $\delta = 1$. The exact solution of (4.6) is

$$u(x,y,t) = \sin(x+y)\exp(-2t), \qquad (x,y,t) \in [0, 2\pi]^2 \times [0, T]. \tag{4.7}$$

Numerical errors at $T = 0.1$ and their convergence rates are listed in Table 4. We observe that our scheme gives the optimal $(k+1)$-th order of the accuracy for Cartesian meshes with Q^k elements, which confirms the result in Theorem 3.2. For Cartesian meshes with P^k elements, we once again observe the $(k+1)$-th order of accuracy. Moreover, we see that

Table 4
Accuracy on (4.6) with $T = 0.1$.

	h	$k=0$		$k=1$		$k=2$	
		e_2	Order	e_2	Order	e_2	Order
Cartesian meshes with P^k elements	$2\pi/3$	1.77E+00	–	6.39E−01	–	1.54E−01	–
	$2\pi/6$	9.21E−01	0.94	1.71E−01	1.90	1.95E−02	2.98
	$2\pi/12$	4.65E−01	0.99	4.36E−02	1.97	2.45E−03	3.00
Cartesian meshes with Q^k elements	$2\pi/3$	1.77E+00	–	5.02E−01	–	4.69E−02	–
	$2\pi/6$	9.21E−01	0.94	1.28E−01	1.97	6.22E−03	2.92
	$2\pi/12$	4.65E−01	0.99	3.21E−02	2.00	7.78E−04	3.00
Triangular meshes with P^k elements	$2\pi/3$	1.71E+00	–	6.44E−01	–	1.05E−01	–
	$2\pi/6$	9.19E−01	0.90	1.71E−01	1.92	1.41E−02	2.90
	$2\pi/12$	4.63E−01	0.99	4.32E−02	1.98	1.77E−03	2.99

Table 5
Accuracy on (4.8) with $T = 0.1$.

	h	$k=0$		$k=1$		$k=2$	
		e_2	Order	e_2	Order	e_2	Order
Cartesian meshes with P^k elements	$2\pi/3$	1.77E+00	–	6.25E−01	–	1.55E−01	–
	$2\pi/6$	9.22E−01	0.94	1.70E−01	1.88	1.95E−02	2.99
	$2\pi/12$	4.65E−01	0.99	4.34E−02	1.97	2.45E−03	3.00
Cartesian meshes with Q^k elements	$2\pi/3$	1.77E+00	–	5.64E−01	–	4.84E−02	–
	$2\pi/6$	9.22E−01	0.94	1.36E−01	2.05	6.35E−03	2.93
	$2\pi/12$	4.65E−01	0.99	3.36E−02	2.02	7.82E−04	3.02
Triangular meshes with P^k elements	$2\pi/3$	1.73E+00	–	6.32E−01	–	1.07E−01	–
	$2\pi/6$	9.21E−01	0.91	1.68E−01	1.92	1.40E−02	2.92
	$2\pi/12$	4.64E−01	0.99	4.20E−02	2.00	1.77E−03	2.99

the scheme with Q^k elements is more accurate than the one with P^k elements since Q^k elements have more degree of freedom. For the triangular meshes, though we only get the sub-optimal order of accuracy result in Theorem 3.2, the numerical experiments indicate that the scheme still has optimal $(k+1)$-th order of accuracy for P^k elements.

4.2.2. Degenerate equation with smooth F in multiple space dimensions

We solve the two-dimensional degenerate fully nonlinear second-order PDE

$$\begin{cases} u_t = \left[\frac{1}{2}(u_{xx})^3 + \frac{1}{2}\left(u_{yy}\right)^3 + u_{xx} + u_{yy} + u_{xy} + u_{yx}\right] + f, \\ u(x,y,0) = \sin(x+y), \end{cases} \tag{4.8}$$

where we take

$$f(x,y,t) := 2\sin(x+y)\exp(-2t) + \sin^3(x+y)\exp(-6t),$$

such that the exact solution of (4.8) is same as the exact solution of (4.6).

We have

$$F_{\mathbb{P}}(\mathbb{P}, \vec{v}, u, x, y, t) = \begin{bmatrix} \frac{3}{2}p_{11}^2 + 1 & 1 \\ 1 & \frac{3}{2}p_{22}^2 + 1 \end{bmatrix} \geq \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \geq 0,$$

which indicates that (4.8) is a degenerate equation.

We check the numerical errors and order of convergence at $T = 0.1$, and we still obtain expected optimal order of accuracy for the multi-dimensional degenerate case, see Table 5.

4.2.3. Degenerate equation with non-smooth F in multiple space dimensions

We approximate the two-dimensional degenerate HJB equation

$$\begin{cases} u_t(x,y,t) &= \min\left\{\frac{1}{2}\left[(x-1)^2 u_{xx} + (y-1)^2 u_{yy}\right], (x-1)^2 u_{xx} + (y-1)^2 u_{yy}\right\} + f(x,y,t), \\ u(x,y,0) &= \sin(x+y), \end{cases} \tag{4.9}$$

where we take

$$f(x,y,t) := \max\left\{\frac{1}{2}\sin(x+y), \sin(x+y)\right\}\left[(x-1)^2 + (y-1)^2\right]\exp(-2t) - 2\sin(x+y)\exp(-2t),$$

such that the exact solution of (4.9) is (4.7).

Table 6
Accuracy on (4.9) with $T = 0.1$.

	h	$k=0$		$k=1$		$k=2$	
		e_2	Order	e_2	Order	e_2	Order
Cartesian meshes with P^k elements	$2\pi/3$	1.82E+00	–	6.23E−01	–	1.53E−01	–
	$2\pi/6$	9.28E−01	0.97	1.65E−01	1.91	1.94E−02	2.98
	$2\pi/12$	4.66E−01	0.99	4.24E−02	1.96	2.44E−03	2.99
Cartesian meshes with Q^k elements	$2\pi/3$	1.82E+00	–	4.59E−01	–	4.62E−02	–
	$2\pi/6$	9.28E−01	0.97	1.18E−01	1.96	5.97E−03	2.95
	$2\pi/12$	4.66E−01	0.99	3.03E−02	1.97	7.63E−04	2.97
Triangular meshes with P^k elements	$2\pi/3$	1.76E+00	–	6.13E−01	–	1.07E−01	–
	$2\pi/6$	9.25E−01	0.92	1.64E−01	1.90	1.40E−02	2.94
	$2\pi/12$	4.64E−01	1.00	4.13E−02	1.99	1.77E−03	2.98

We have

$$F_{\mathbb{P}}(\mathbb{P}, \vec{v}, u, x, y, t) = \frac{1}{2}\left(1 + 1_{\left\{(x-1)^2 p_{11} + (y-1)^2 p_{22} \le 0\right\}}\right)\begin{bmatrix}(x-1)^2 & 0\\ 0 & (y-1)^2\end{bmatrix} \ge 0,$$

which indicates that (4.9) is a degenerate equation.

Both the errors and numerical orders of accuracy are listed in Table 6. We once again observe the designed $(k+1)$-th order for this degenerate PDE with non-smooth F.

5. Applications to stochastic optimal control

We now show that one can solve stochastic optimal control problems by using our LDG method. To this end, let us consider the stochastic control problem. The random effect is inherent in most real-world systems. It places many disadvantages (and sometimes, surprisingly, advantages) on humankind's efforts, which are usually associated with the quest for optimal results. The dynamic state equation is described by a controlled stochastic differential equation (SDE)

$$\begin{cases} dX^{\alpha(\cdot)}(t) &= b(t, X^{\alpha(\cdot)}(t), \alpha(t))\,dt + \sigma(t, X^{\alpha(\cdot)}(t), \alpha(t))\,dW(t), \quad t \in (s, T],\\ X^{\alpha(\cdot)}(s) &= y, \end{cases}$$

with the cost functional

$$J(s, y; \alpha(\cdot)) = \mathbb{E}\left[\int_s^T f(t, X^{\alpha(\cdot)}(t), \alpha(t))dt + h(X^{\alpha(\cdot)}(T))\right],$$

where $\left(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\ge 0}, \mathscr{P}\right)$ is a complete filtered probability space on which a d'-dimensional Wiener process $W = \{W(t);\ t \ge 0\}$ is defined such that $\{\mathscr{F}_t\}_{t\ge 0}$ is the natural filtration generated by W, augmented by all the $\mathscr{P}$-null sets in $\mathscr{F}$. The basic source of uncertainty in diffusion models is white noise, which represents the joint effects of a large number of independent random forces acting on the systems. The decision makers must select an optimal decision among all possible ones to achieve the best expected result related to their goals in the stochastic environment. We aim to find the optimal control $\alpha^*(\cdot)$ from the feasible control set

$$\mathcal{U}[s, T] := \left\{\alpha : [s, T] \times \Omega \longrightarrow U \mid \alpha(\cdot) \text{ is } \{\mathscr{F}_t\}_{s\le t\le T}\text{-adapted}\right\},$$

such that

$$J(s, y; \alpha^*(\cdot)) = \min_{\alpha(\cdot)\in\mathcal{U}[s,T]} J(s, y; \alpha(\cdot)).$$

We define

$$V(s, y) := \min_{\alpha(\cdot)\in\mathcal{U}[s,T]} J(s, y; \alpha(\cdot)),$$

which is called the value function. The above formulations of the stochastic optimal control can have several concrete application in the real world such as *production planning*, *investment vs. consumption*, *reinsurance and dividend management*, *technology diffusion*, *queueing systems in heavy traffic*. For more details, we refer to [4] and the references therein.

According to the stochastic verification theorem in the dynamic programming (see e.g. [4,27]), the value function V is determined by the backward HJB equation involving the second-order spatial derivative

$$\begin{cases} -V_t(x, t) &= \min_{\alpha\in U} G\left(D^2V(x, t), \nabla V(x, t), x, t, \alpha\right), \qquad (x, t) \in \mathbb{R}^n \times [0, T),\\ V(x, T) &= h(x), \end{cases}$$

where

$$G(\mathbb{P}, \vec{v}, x, t, \alpha) = \frac{1}{2} tr\left(\mathbb{P}\sigma(t, x, \alpha)\sigma(t, x, \alpha)^T\right) + \langle \vec{v}, b(t, x, \alpha)\rangle + f(t, x, \alpha).$$

Moreover, for each $(x, t) \in \mathbb{R}^n \times [0, T]$, we have

$$\overline{\alpha}(x, t) := \mathbb{E}\left[\alpha^*(t) \,\middle|\, X^{\alpha^*(\cdot)}(t) = x\right] = \arg\min_{\alpha \in U} G\left(D^2 V(x, t), \nabla V(x, t), x, t, \alpha\right). \tag{5.1}$$

Thus the key is to solve HJB equation. Now, we consider the following one-dimensional ($d = d' = 1$) linear system

$$\begin{cases} dX^{\alpha(\cdot)}(t) &= \left[\beta_1 X^{\alpha(\cdot)}(t) + \beta_2 \alpha(t) + \beta_3\right] dt + \left[\sigma_1 X^{\alpha(\cdot)}(t) + \sigma_2 \alpha(t) + \sigma_3\right] dW(t), \quad t \in (s, T], \\ X^{\alpha(\cdot)}(s) &= y, \end{cases}$$

with the $\alpha(\cdot)$-quadratic cost functional

$$J(s, y; \alpha(\cdot)) = \mathbb{E}\left[\int_s^T f(t, X^{\alpha(\cdot)}(t))dt + q\int_s^T |\alpha(t)|^2\, dt + \sin(X^{\alpha(\cdot)}(T))\right],$$

where the function f is left to determined later, and the parameters $q \in \mathbb{R}_+$, $\vec{\beta} = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$, $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3) \in \mathbb{R}^3$ are given. One can construct the associate backward HJB equation

$$-V_t = \min_{\alpha \in U}\left\{\frac{(\sigma_1 x + \sigma_2 \alpha + \sigma_3)^2}{2} V_{xx} + (\beta_1 x + \beta_2 \alpha + \beta_3) V_x + q\alpha^2\right\} + f(t, x),$$

with $V(x, T) = \sin(x)$. We define $u(x, t) := V(x, T - t)$. Then the above HJB equation reads

$$u_t = \min_{\alpha \in U}\left\{\left(\frac{\sigma_2^2}{2} u_{xx} + q\right)\alpha^2 + (\sigma_1\sigma_2 x u_{xx} + \sigma_2\sigma_3 u_{xx} + \beta_2 u_x)\alpha\right\} + \frac{1}{2}(\sigma_1 x + \sigma_3)^2 u_{xx} + (\beta_1 x + \beta_3) u_x + f(t, x), \tag{5.2}$$

with $u(x, 0) = \sin(x)$.

5.1. Continuous case

Let $U := \mathbb{R}$. In (5.2), we choose $q > 0$ large enough such that $\frac{\sigma_2^2}{2} u_{xx} + q > 0$. According to (5.1), the optimal control is

$$\overline{\alpha}(x, t) = -\frac{(\sigma_1\sigma_2 x + \sigma_2\sigma_3) u_{xx}(x, T - t) + \beta_2 u_x(x, T - t)}{\sigma_2^2 u_{xx}(x, T - t) + 2q}, \tag{5.3}$$

and the corresponding HJB equation is

$$u_t = -\frac{1}{2}\frac{(\sigma_1\sigma_2 x u_{xx} + \sigma_2\sigma_3 u_{xx} + \beta_2 u_x)^2}{\sigma_2^2 u_{xx} + 2q} + \frac{1}{2}(\sigma_1 x + \sigma_3)^2 u_{xx} + (\beta_1 x + \beta_3) u_x + f(t, x), \tag{5.4}$$

with $u(x, 0) = \sin(x)$, where we take

$$f(x, t) := e^{-t}\left[\frac{1}{2}(\sigma_1 x + \sigma_3)^2 \sin(x) - (\beta_1 x + \beta_3)\cos(x) - \sin(x)\right] + \frac{e^{-2t}\left[\beta_2\cos(x) - \sigma_1\sigma_2 x\sin(x) - \sigma_2\sigma_3\sin(x)\right]^2}{4q - 2\sigma_2^2 e^{-t}\sin(x)},$$

such that the exact solution of (5.4) is (4.3).

We use the LDG method (3.6) to solve (5.4), and approximate the optimal control (5.3) by

$$\overline{\alpha}_h(x, t) = -\frac{(\sigma_1\sigma_2 x + \sigma_2\sigma_3) p_h(x, T - t) + \beta_2 v_h(x, T - t)}{\sigma_2^2 p_h(x, T - t) + 2q}. \tag{5.5}$$

In the numerical test, we set $T = 0.1$, $q = 15$, $\vec{\beta} = \left(\frac{1}{5}, \frac{1}{2}, \frac{1}{5}\right)$, $\vec{\sigma} = \left(\frac{1}{5}, \frac{1}{5}, 1\right)$. We compute L^2-errors for the HJB solution and the optimal control

$$e_2 := \|u_h(\cdot, T) - u(\cdot, T)\|, \qquad e_2^\alpha := \|\overline{\alpha}_h(\cdot, 0) - \overline{\alpha}(\cdot, 0)\|.$$

The numerical results for e_2 and e_2^α are shown in Tables 7 and 8. It can be seen that our LDG scheme has optimal $(k+1)$-th order of accuracy both for the HJB solution u and the optimal control $\overline{\alpha}(\cdot)$.

5.2. Discontinuous case

In (5.2), we choose

$$U = [0, 1], \quad \sigma_2 = 1, \quad q = \beta_1 = \beta_2 = \beta_3 = \sigma_1 = \sigma_3 = 0,$$

Table 7
Accuracy on (5.4) with $T = 0.1$, $q = 15$, $\vec{\beta} = (\frac{1}{5}, \frac{1}{2}, \frac{1}{5})$, $\vec{\sigma} = (\frac{1}{5}, \frac{1}{5}, 1)$.

h	$k=0$		$k=1$		$k=2$		$k=3$	
	e_2	Order	e_2	Order	e_2	Order	e_2	Order
$2\pi/10$	2.94E−01	–	3.87E−02	–	1.94E−03	–	7.48E−05	–
$2\pi/20$	1.47E−01	1.00	9.65E−03	2.00	2.43E−04	3.00	4.69E−06	4.00
$2\pi/40$	7.36E−02	1.00	2.41E−03	2.00	3.03E−05	3.00	2.93E−07	4.00

Table 8
Accuracy on (5.3), (5.5) with $T = 0.1$, $q = 15$, $\vec{\beta} = (\frac{1}{5}, \frac{1}{2}, \frac{1}{5})$, $\vec{\sigma} = (\frac{1}{5}, \frac{1}{5}, 1)$.

h	$k=0$		$k=1$		$k=2$		$k=3$	
	e_2^α	Order	e_2^α	Order	e_2^α	Order	e_2^α	Order
$2\pi/10$	1.02E−02	–	6.44E−04	–	3.09E−05	–	1.28E−06	–
$2\pi/20$	5.02E−03	1.02	1.63E−04	1.98	3.86E−06	3.00	8.11E−08	3.98
$2\pi/40$	2.49E−03	1.01	4.08E−05	2.00	4.83E−07	3.00	5.08E−09	4.00

Table 9
Accuracy on (5.6) with $T = 0.1$.

h	$k=0$		$k=1$		$k=2$		$k=3$	
	e_2	Order	e_2	Order	e_2	Order	e_2	Order
$2\pi/10$	2.89E−01	–	3.46E−02	–	1.71E−03	–	6.39E−05	–
$2\pi/20$	1.45E−01	0.99	9.09E−03	1.93	2.15E−04	2.99	3.99E−06	4.00
$2\pi/40$	7.27E−02	1.00	2.30E−03	1.98	2.66E−05	3.01	2.49E−07	4.00

Table 10
Accuracy on (5.7) and (5.8) with $T = 0.1$.

h	$k=0$		$k=1$		$k=2$		$k=3$	
	e_2^α	Order	e_2^α	Order	e_2^α	Order	e_2^α	Order
$2\pi/10$	5.61E−02	–	7.15E−01	–	1.63E−01	–	9.71E−02	–
$2\pi/20$	3.96E−02	0.50	2.49E−01	1.52	7.93E−02	1.04	3.96E−02	1.29
$2\pi/40$	2.80E−02	0.50	2.80E−02	3.15	3.96E−02	1.00	2.80E−02	0.50

and

$$f(x,t) = -\frac{e^{-t}}{2}\left(1 + 1_{\{\sin(x)<0\}}\right)\sin(x).$$

Then the corresponding HJB equation is

$$u_t = \frac{1}{2}\min\{0\,,u_{xx}\} - \frac{e^{-t}}{2}\left(1 + 1_{\{\sin(x)<0\}}\right)\sin(x), \tag{5.6}$$

with $u(x,0) = \sin(x)$; and the optimal control is

$$\overline{\alpha}(x,t) = 1 - 1_{\{u_{xx}(x,T-t)\geq 0\}}. \tag{5.7}$$

The exact solution of (5.6) is (4.3). We use the LDG method (3.6) to solve (5.6), and approximate the optimal control (5.7) by

$$\overline{\alpha}_h(x,t) = 1 - 1_{\{p_h(x,T-t)\geq 0\}}. \tag{5.8}$$

We set $T = 0.1$. The numerical results are shown in Tables 9 and 10. It can seen from Table 9 that our LDG scheme has optimal $(k+1)$-th order of accuracy for the HJB solution u. Since the exact optimal control (5.7) is discontinuous, we do not expect $(k+1)$-th order of accuracy in Table 10. However, Fig. 1 shows that our scheme works very well, which not only approximates accurately the optimal control in the smooth region, but also resolves the discontinuity sharply. In general, many applications will not satisfy the assumptions in Theorem 3.1 and Theorem 3.2. Nevertheless, the numerical results in this section indicates that our scheme still works well in some more general frameworks.

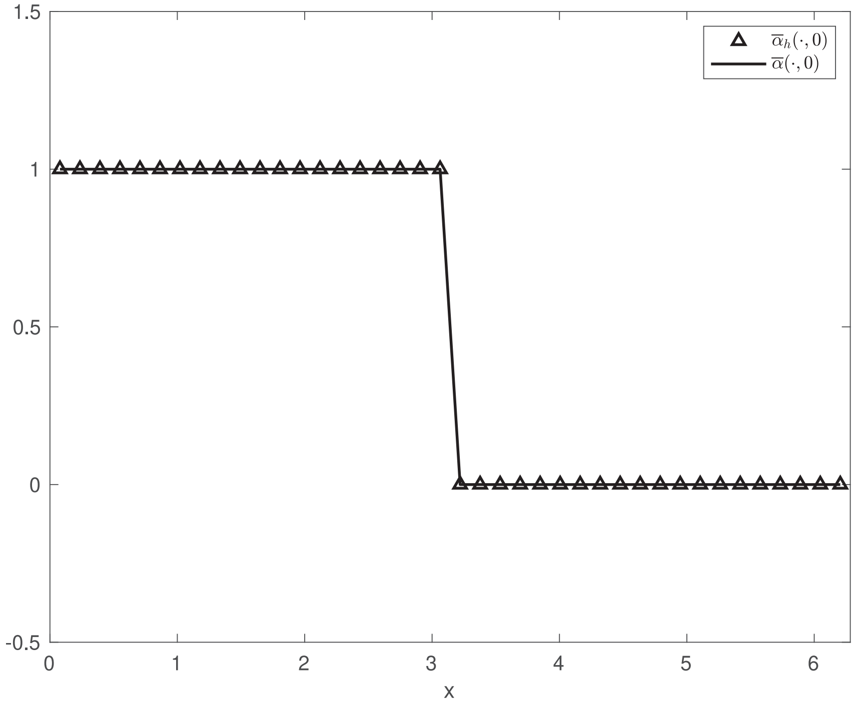

Fig. 1. Test on (5.7) and (5.8) with $T = 0.1$, $k = 3$, $N = 40$.

6. Proofs

6.1. Preliminaries

6.1.1. Projection properties

For one-dimensional case, we consider the standard L^2-projection (denoted by $\mathcal{P}$), and the local Gauss–Radau projections (denoted by $\mathcal{P}^{\pm}$) into space V_h^k. For each $j = 1, \ldots, N$, the projections satisfy that

$$\int_{I_j} [\mathcal{P}f(x) - f(x)]\, g(x)\, dx = 0, \quad \forall g \in P^k(I_j);$$

$$\int_{I_j} \left[\mathcal{P}^+ f(x) - f(x)\right] g(x)\, dx = 0, \quad \forall g \in P^{k-1}(I_j), \quad \text{and} \quad \mathcal{P}^+ f(x^+_{j-\frac{1}{2}}) = f(x_{j-\frac{1}{2}});$$

$$\int_{I_j} \left[\mathcal{P}^- f(x) - f(x)\right] g(x)\, dx = 0, \quad \forall g \in P^{k-1}(I_j), \quad \text{and} \quad \mathcal{P}^- f(x^-_{j+\frac{1}{2}}) = f(x_{j+\frac{1}{2}}).$$

For two-dimensional case, to prove the sub-optimal error estimates with triangular meshes, we are going to use the standard L^2-projection, which is denoted by Π. To prove the optimal error estimates for two-dimensional problems with Cartesian meshes, we need suitable projections $\Pi^{\pm}$ similar to the one-dimensional case. We set $\vec{\beta} := (1, 1)$ and use projections in [28,29]. Define $\Pi^{\pm} := \mathcal{P}_x^{\pm} \otimes \mathcal{P}_y^{\pm}$, where the subscripts indicate the application of the one-dimensional operators $\mathcal{P}^{\pm}$ with respect to the corresponding variable. Denote the rectangular partition by $I_j \times J_i$, where $I_j = \left[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right]$ and $J_i = \left[y_{i-\frac{1}{2}}, y_{i+\frac{1}{2}}\right]$. For each $j = 1, \ldots, N_x$ and $i = 1, \ldots, N_y$, the projections $\Pi^{\pm}$ satisfy:

$$\int_{I_j}\int_{J_i} \left(\Pi^{\pm} f(x, y) - f(x, y)\right) g(x, y) dy dx = 0, \tag{6.1a}$$

for any $g \in \left[P^{k-1}(I_j) \otimes P^k(J_i)\right] \cup \left[P^k(I_j) \otimes P^{k-1}(J_i)\right]$; and

$$\int_{J_m} \left(\Pi^+ f(x^+_{n-\frac{1}{2}}, y) - f(x_{n-\frac{1}{2}}, y)\right) g(x^+_{n-\frac{1}{2}}, y) dy = 0, \tag{6.1b}$$

$$\int_{I_n} \left(\Pi^+ f(x, y^+_{m-\frac{1}{2}}) - f(x, y_{m-\frac{1}{2}})\right) g(x, y^+_{m-\frac{1}{2}}) dx = 0, \tag{6.1c}$$

$$\int_{J_m}\left(\Pi^- f(x^-_{n+\frac{1}{2}}, y) - f(x_{n+\frac{1}{2}}, y)\right) g(x^-_{n+\frac{1}{2}}, y)dy = 0, \tag{6.1d}$$

$$\int_{I_n}\left(\Pi^- f(x, y^-_{m+\frac{1}{2}}) - f(x, y_{m+\frac{1}{2}})\right) g(x, y^-_{m+\frac{1}{2}})dx = 0, \tag{6.1e}$$

for any $g \in Q^k(I_n \otimes J_m)$.

Let $\mathcal{Q}$ be the projections $\mathcal{P}$, $\mathcal{P}^{\pm}$, Π and $\Pi^{\pm}$. Then for any function f in H^{k+1}, we have (c.f. [28–31])

$$\|\mathcal{Q}f - f\| + h\,\|\nabla(\mathcal{Q}f - f)\| + h^{\frac{1}{2}}\,\|\mathcal{Q}f - f\|_{\Gamma_h} \le Ch^{k+1}, \tag{6.2}$$

where the positive constant C is independent of h.

6.1.2. Inverse properties

Finally, we list some inverse properties of the finite element space V_h^k. For any function w_h in V_h^k, there exists a positive constant C, independent of h, such that

$$h\,\|\nabla w_h\| + h^{\frac{1}{2}}\,\|w_h\|_{\Gamma_h} + h^{\frac{d}{2}}\,\|w_h\|_\infty \le C\,\|w_h\|\,, \tag{6.3}$$

where d is the spatial dimension. More details of the inverse properties can be found in [31].

6.2. Proof of Theorem 3.1

Proof. For $Z = u, v, p$, we define

$$e_Z := Z - Z_h = \xi_Z - \eta_Z, \quad \text{with} \quad \xi_Z := \mathcal{P}^Z Z - Z_h, \quad \eta_Z := \mathcal{P}^Z Z - Z,$$

where we choose $\mathcal{P}^u := \mathcal{P}^-$, $\mathcal{P}^v := \mathcal{P}^+$ and $\mathcal{P}^p := \mathcal{P}$.

We denote

$$\Theta := (p, v, u, \cdot)\,(x, t) \quad \text{and} \quad \Theta_h := (p_h, v_h, u_h, \cdot)\,(x, t)\,.$$

By taking a Taylor expansion, we have

$$\begin{aligned} F(\Theta) - F(\Theta_h) &= F(p, v, u, x, t) - F(p_h, v, u, x, t) + F(p_h, v, u, x, t) \\ &\quad - F(p_h, v_h, u, x, t) + F(p_h, v_h, u, x, t) - F(p_h, v_h, u_h, x, t) \\ &= F_p\,(\Theta)\,e_p - \frac{1}{2}F_{pp}\left(\overline{\Theta}_p\right)|e_p|^2 + F_v\left(\overline{\Theta}_v\right)e_v + F_u\left(\overline{\Theta}_u\right)e_u. \end{aligned}$$

Then the cell error equation is

$$\begin{cases} \displaystyle\int_{I_j} (e_u)_t z_{h,u} dx = \int_{I_j}\left\{F_p\,(\Theta)\,e_p - \frac{1}{2}F_{pp}\left(\overline{\Theta}_p\right)\left|e_p\right|^2 + F_v\left(\overline{\Theta}_v\right)e_v + F_u\left(\overline{\Theta}_u\right)e_u\right\} z_{h,u} dx, & \text{(a)} \\ \displaystyle\int_{I_j} e_v z_{h,v} dx = H_j^-\big(e_u\,,\,z_{h,v}\big), & \text{(b)} \\ \displaystyle\int_{I_j} e_p z_{h,p} dx = H_j^+\big(e_v\,,\,z_{h,p}\big). & \text{(c)} \end{cases} \tag{6.4}$$

Since the projections in this paper are performed only for the spatial variable, the defined functions e_u, ξ_u, η_u admit time derivative according to the regularity of u.

Define $\alpha_j(t) := F_p(\Theta)(x_j, t)$. Taking $z_{h,u} = \xi_u$ in (6.4a), $z_{h,v} = \alpha_j\xi_v$ in (6.4b), $z_{h,p} = \alpha_j\xi_u$ in (6.4c), adding the resulting equations and summing over j from 1 to N, we have

$$\int_D (\xi_u)_t\,\xi_u dx + \sum_{j=1}^N \alpha_j \int_{I_j} |\xi_v|^2\,dx = \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_4,$$

with

$$\mathcal{T}_1 = \sum_{j=1}^N \int_{I_j}\left\{F_p\,(\Theta) - \alpha_j\right\} e_p \xi_u dx - \frac{1}{2}\int_D F_{pp}\left(\overline{\Theta}_p\right)\left|e_p\right|^2 \xi_u dx,$$

$$\mathcal{T}_2 = \sum_{j=1}^N \alpha_j\left\{H_j^-\left(\xi_u\,,\,\xi_v\right) + H_j^+\left(\xi_v\,,\,\xi_u\right)\right\},$$

$$\mathcal{T}_3 = -\sum_{j=1}^{N} \alpha_j \left\{ H_j^- \left(\eta_u \,,\, \xi_v \right) + H_j^+ \left(\eta_v \,,\, \xi_u \right) \right\},$$

$$\mathcal{T}_4 = \int_D \left\{ (\eta_u)_t + F_v \left(\overline{\Theta}_v \right) e_v + F_u \left(\overline{\Theta}_u \right) e_u \right\} \xi_u dx + \sum_{j=1}^{N} \alpha_j \int_{I_j} \eta_v \xi_v dx,$$

where the terms $\mathcal{T}_i$, $i = 1, 2, 3, 4$ will be estimated separately later.

• **The estimate of $\mathcal{T}_1$.** To deal with the nonlinearity of the function F, we we treat it by Taylor expansion and make a priori assumption that for sufficiently small h, there holds

$$\|e_u\| \le \mathbb{C}h^3. \tag{6.5}$$

This a priori assumption is frequently used in the analysis of nonlinear problems (see e.g. [17,23–25]), and the reasonableness of this a priori assumption will be justified later. In fact, if the flux function F is linear with respect to the first and the second variables, i.e. $F(p, v, u, x, t) = C_1 p + C_2 v + f(u, x, t)$ for some constants C_1 and C_2, one can check that this a priori assumption is not necessary and can be removed. According to the inverse inequality (6.3) and the projection properties (6.2) with $k \ge 2$, we get

$$\|\xi_u\|_\infty \le Ch^{-\frac{1}{2}} \|\xi_u\| \le Ch^{-\frac{1}{2}} \left(\|e_u\| + \|\eta_u\| \right) \le Ch^{\frac{5}{2}}.$$

Since u and F are smooth enough, we have

$$\sup_{1 \le j \le N} \left\| F_p \left(\Theta \right) - \alpha_j \right\|_{L^\infty(I_j \times (0,T))} \le Ch.$$

It follows that

$$|\mathcal{T}_1| \le Ch \left\| e_p \right\| \|\xi_u\| + Ch^{\frac{5}{2}} \left\| e_p \right\|^2.$$

Next we estimate $\left\| e_p \right\|$. Define $q_h := (v_h)_x - p_h$. Note that

$$\int_{I_j} (v_h)_x \, q_h dx = -\int_{I_j} v_h \left(q_h \right)_x dx + v_h \left(x_{j+\frac{1}{2}}^- \right) q_h \left(x_{j+\frac{1}{2}}^- \right) - v_h \left(x_{j-\frac{1}{2}}^+ \right) q_h \left(x_{j-\frac{1}{2}}^+ \right).$$

Taking $z_{h,p} = q_h$ in (3.6c), we have

$$\int_{I_j} |q_h|^2 \, dx = - \left[v_h \right]_{j+\frac{1}{2}} q_{h,j+\frac{1}{2}}^-.$$

Summing over j, we get

$$\|q_h\|^2 = -\sum_{j=1}^{N} \left[v_h \right]_{j+\frac{1}{2}} q_{h,j+\frac{1}{2}}^- = \sum_{j=1}^{N} \left[v - v_h \right]_{j+\frac{1}{2}} q_{h,j+\frac{1}{2}}^- \le C \|\xi_v - \eta_v\|_{\Gamma_h} \|q_h\|_{\Gamma_h} \le Ch^{-1} \left(\|\xi_v\| + h^{k+1} \right) \|q_h\|.$$

Thus we have

$$\|q_h\| \le Ch^{-1} \left(\|\xi_v\| + h^{k+1} \right).$$

Since

$$e_p = p - p_h = (v - v_h)_x + (v_h)_x - p_h = (\xi_v)_x - (\eta_v)_x + q_h,$$

it yields that

$$\left\| e_p \right\| \le Ch^{-1} \left(\|\xi_v\| + h^{k+1} \right).$$

By Young inequality, for small enough h, we have

$$|\mathcal{T}_1| \le \frac{\delta}{4} \|\xi_v\|^2 + C \|\xi_u\|^2 + Ch^{2k+2}.$$

• **The estimate of $\mathcal{T}_2$.** For any $f, g \in V_h^k$, we have

$$H_j^+(f, g) + H_j^-(g, f) = -f_{j+\frac{1}{2}}^- g_{j+\frac{1}{2}}^- + f_{j-\frac{1}{2}}^+ g_{j-\frac{1}{2}}^+ + f_{j+\frac{1}{2}}^+ g_{j+\frac{1}{2}}^- - f_{j-\frac{1}{2}}^+ g_{j-\frac{1}{2}}^+ + g_{j+\frac{1}{2}}^- f_{j+\frac{1}{2}}^- - g_{j-\frac{1}{2}}^- f_{j-\frac{1}{2}}^+ = f_{j+\frac{1}{2}}^+ g_{j+\frac{1}{2}}^- - f_{j-\frac{1}{2}}^+ g_{j-\frac{1}{2}}^-.$$

According to the periodicity, we get

$$\mathcal{T}_2 = \sum_{j=1}^{N} \alpha_j \left(\xi_{v,j+\frac{1}{2}}^+ \xi_{u,j+\frac{1}{2}}^- - \xi_{v,j-\frac{1}{2}}^+ \xi_{u,j-\frac{1}{2}}^- \right) = \sum_{j=1}^{N} \left(\alpha_j - \alpha_{j+1} \right) \xi_{v,j+\frac{1}{2}}^+ \xi_{u,j+\frac{1}{2}}^-.$$

Note that

$$\sup_{1\le j\le N}\left|\alpha_j-\alpha_{j+1}\right|=\sup_{1\le j\le N}\left|F_p(\Theta)(x_j,t)-F_p(\Theta)(x_{j+1},t)\right|\le Ch.$$

It follows that

$$|\mathcal{T}_2|\le Ch\left(\sum_{j=1}^{N}\left|\xi^{+}_{v,j+\frac12}\right|^2\right)^{\frac12}\left(\sum_{j=1}^{N}\left|\xi^{-}_{u,j+\frac12}\right|^2\right)^{\frac12}\le Ch\left\|\xi_v\right\|_{\Gamma_h}\left\|\xi_u\right\|_{\Gamma_h}\le C\left\|\xi_v\right\|\left\|\xi_u\right\|\le\frac{\delta}{4}\left\|\xi_v\right\|^2+C\left\|\xi_u\right\|^2.$$

• **The estimate of $\mathcal{T}_3$.** By the definition of projections $\mathcal{P}^{\pm}$ (see Section 6.1.1), we have, for any $j=1,2,\ldots,N$, and $z_{h,u},z_{h,p}\in V_h^k$,

$$H_j^-\left(\eta_u(\cdot,t),z_{h,u}\right)=0\quad\text{and}\quad H_j^+\left(\eta_p(\cdot,t),z_{h,p}\right)=0.$$

Since $\xi_u,\xi_v\in V_h^k$, we have $\mathcal{T}_3\equiv 0$.
• **The estimate of $\mathcal{T}_4$.** Since u and F are smooth enough with bounded derivatives, we have

$$|\mathcal{T}_4|\le\frac{\delta}{4}\left\|\xi_v\right\|^2+C\left\|\xi_u\right\|^2+Ch^{2k+2}.$$

Concluding above, we get

$$\int_D(\xi_u)_t\,\xi_u dx+\sum_{j=1}^{N}\alpha_j\int_{I_j}|\xi_v|^2\,dx\le\frac{3\delta}{4}\left\|\xi_v\right\|^2+C\left\|\xi_u\right\|^2+Ch^{2k+2}.$$

Since F is uniformly elliptic, i.e., $F_p(\Theta)\ge\delta$, we have $\alpha_j\ge\delta$ for each $j=1,\ldots,N$. It follows that

$$\frac12\frac{d}{dt}\left\|\xi_u\right\|^2+\frac{\delta}{4}\left\|\xi_v\right\|^2\le C\left\|\xi_u\right\|^2+Ch^{2k+2}.$$

Then the Gronwall's inequality tells us that

$$\sup_{0\le t\le T}\left\|\xi_u(\cdot,t)\right\|^2+\int_0^T\left\|\xi_v(\cdot,t)\right\|^2dt\le Ch^{2k+2}.$$

According to the triangle inequality, we finally get the optimal error estimate for the LDG scheme

$$\sup_{0\le t\le T}\left\|u(\cdot,t)-u_h(\cdot,t)\right\|+\left(\int_0^T\left\|u_x(\cdot,t)-v_h(\cdot,t)\right\|^2dt\right)^{\frac12}\le Ch^{k+1}.\tag{6.6}$$

Now, to complete the proof, let us follow [17,23–25] to verify the a priori assumption (6.5). In fact, the inequality (6.6) for $k\ge 2$ implies that the a priori assumption (6.5) is true. □

6.3. Proof of Theorem 3.2

Proof. We define $\vec{e}_v=(e_v^1,\ldots,e_v^d)^{\mathrm{T}}$ with

$$e_v^l:=v^l-v_h^l=\xi_v^l-\eta_v^l,\quad \xi_v^l:=\Pi^v v^l-v_h^l,\quad \eta_v^l:=\Pi^v v^l-v^l,\qquad l=1,\ldots,d.$$

Also, ξ_u and η_u are defined in a similar way, and we denote

$$e_{\mathbb{P}}:=\left[e_{\mathbb{P}}^{ml}\right]_{d\times d}=\left[p^{ml}-p_h^{ml}\right]_{d\times d}.$$

The projections Π^v and Π^u are chosen from $\left\{\Pi,\Pi^{\pm}\right\}$ defined in Section 6.1.1, which will be determined later.
We denote

$$\Theta:=(\mathbb{P},\vec{v},u,\cdot)(x,t)\quad\text{and}\quad\Theta_h:=(\mathbb{P}_h,\vec{v}_h,u_h,\cdot)(x,t).$$

The cell error equation is

$$\begin{cases}\displaystyle\int_K(e_u)_t\,z_{h,u}dx=\int_K\{F(\Theta)-F(\Theta_h)\}\,z_{h,u}\,dx, & \text{(a)}\\ \displaystyle\int_K\vec{e}_v\cdot\vec{z}_{h,v}dx=H_K^-\left(e_u,\vec{z}_{h,v}\right), & \text{(b)}\\ \displaystyle\int_K\sum_{m,l=1}^{d}A_K^{ml}e_{\mathbb{P}}^{ml}z_{h,p}dx=H_K^+\left(A_K\vec{e}_v,z_{h,p}\right). & \text{(c)}\end{cases}\tag{6.7}$$

By taking a Taylor expansion, we have

$$\begin{aligned} F(\Theta)-F(\Theta_h) &= F(\mathbb{P},\vec{v},u,x,t)-F(\mathbb{P}_h,\vec{v},u,x,t)+F(\mathbb{P}_h,\vec{v},u,x,t)\\ &\quad -F(\mathbb{P}_h,\vec{v}_h,u,x,t)+F(\mathbb{P}_h,\vec{v}_h,u,x,t)-F(\mathbb{P}_h,\vec{v}_h,u_h,x,t)\\ &= \sum_{m,l=1}^{d} e_{\mathbb{P}}^{ml}\frac{\partial F}{\partial p^{ml}}(\Theta)-\frac{1}{2}\left(\sum_{m,l=1}^{d} e_{\mathbb{P}}^{ml}\frac{\partial}{\partial p^{ml}}\right)^2 F\left(\overline{\Theta}_{\mathbb{P}}\right)+F_{\vec{v}}\left(\overline{\Theta}_v\right)\cdot\vec{e}_v+F_u\left(\overline{\Theta}_u\right)e_u. \end{aligned}$$

We define

$$\Lambda_K(t)=\left[\Lambda_K^{ml}\right]_{d\times d}(t):=\left[\frac{\partial F}{\partial p^{ml}}(\Theta)(x_K,t)\right]_{d\times d}.$$

Taking $z_{h,u}=\xi_u$ in (6.7a), $\vec{z}_{h,v}=\Lambda_K\vec{\xi}_v$ in (6.7b), $A_K=\Lambda_K$ and $z_{h,p}=\xi_u$ in (6.7c), adding the resulting equations and summing over K, we have

$$\int_D(\xi_u)_t\,\xi_u dx+\sum_{K\in\mathcal{T}_h}\int_K\vec{\xi}_v^{\,\mathrm{T}}\Lambda_K\vec{\xi}_v dx=\mathcal{T}_1+\mathcal{T}_2+\mathcal{T}_3+\mathcal{T}_4,$$

with

$$\begin{aligned} \mathcal{T}_1 &= \sum_{K\in\mathcal{T}_h}\int_K\left\{\sum_{m,l=1}^{d} e_{\mathbb{P}}^{ml}\left(\frac{\partial F}{\partial p^{ml}}(\Theta)-\Lambda_K^{ml}\right)\right\}\xi_u dx\\ &-\frac{1}{2}\int_D\left(\sum_{m,l=1}^{d} e_{\mathbb{P}}^{ml}\frac{\partial}{\partial p^{ml}}\right)^2 F\left(\overline{\Theta}_{\mathbb{P}}\right)\xi_u dx,\\ \mathcal{T}_2 &= \sum_{K\in\mathcal{T}_h}\left\{H_K^-\left(\xi_u\,,\ \Lambda_K\vec{\xi}_v\right)+H_K^+\left(\Lambda_K\vec{\xi}_v\,,\ \xi_u\right)\right\},\\ \mathcal{T}_3 &= -\sum_{K\in\mathcal{T}_h}\left\{H_K^-\left(\eta_u\,,\ \Lambda_K\vec{\xi}_v\right)+H_K^+\left(\Lambda_K\vec{\eta}_v\,,\ \xi_u\right)\right\},\\ \mathcal{T}_4 &= \int_D\left\{(\eta_u)_t+F_{\vec{v}}\left(\overline{\Theta}_v\right)\cdot\vec{e}_v+F_u\left(\overline{\Theta}_u\right)e_u\right\}\xi_u dx+\sum_{K\in\mathcal{T}_h}\int_K\vec{\eta}_v^{\,\mathrm{T}}\Lambda_K\vec{\xi}_v dx, \end{aligned}$$

where the terms $\mathcal{T}_i$, $i=1,2,3,4$ will be estimated separately later.

• **The estimate of $\mathcal{T}_1$.** To deal with the nonlinearity of the function F, we we treat it by Taylor expansion and make a priori assumption that for sufficiently small h, there holds

$$\|e_u\|\le\mathbb{C}h^4. \tag{6.8}$$

This a priori assumption is frequently used in the analysis of nonlinear problems (see e.g. [17,23–25]), and the reasonableness of this a priori assumption will be justified later. Similar to the one dimensional case, if the flux function F is linear with respect to the first and the second variables, this a priori assumption is not necessary and can be removed. According to the inverse inequality (6.3) and the projection properties (6.2) with $k\ge 3$, we get

$$\|\xi_u\|_\infty\le Ch^{-\frac{d}{2}}\|\xi_u\|\le Ch^{-1}\left(\|e_u\|+\|\eta_u\|\right)\le Ch^3.$$

Since u and F are smooth enough, we have

$$\sup_{K\in\mathcal{T}_h}\left\|\frac{\partial F}{\partial p^{ml}}(\Theta)-\Lambda_K^{ml}\right\|_{L^\infty(K\times(0,T))}\le Ch,$$

for any $m,l=1,\ldots,d$,. It follows that

$$|\mathcal{T}_1|\le Ch\|e_{\mathbb{P}}\|\,\|\xi_u\|+Ch^3\|e_{\mathbb{P}}\|^2. \tag{6.9}$$

Next we estimate $\|e_{\mathbb{P}}\|$. Define $\mathbb{Q}_h:=\left[q_h^{ml}\right]_{d\times d}$ with $q_h^{ml}=\frac{\partial v_h^l}{\partial x^m}-p_h^{ml}$. For any $A_K\in\mathbb{R}^{d\times d}$ and $z_{h,p}\in V_h^k$, we have

$$\int_K\sum_{m,l=1}^{d}A_K^{ml}\frac{\partial v_h^l}{\partial x^m}z_{h,p}dx=-\int_K A_K\vec{v}_h\cdot\nabla z_{h,p}dx+\left(z_{h,p}|^{intK},A_K\vec{v}_h|^{intK}\cdot\vec{n}_K\right)_{\partial K}.$$

It yields that

$$\int_K\sum_{m,l=1}^{d}A_K^{ml}q_h^{ml}z_{h,p}dx=\left(z_{h,p}|^{intK},A_K\left(\vec{v}_h|^{intK}-\vec{v}_h^{+}\right)\cdot\vec{n}_K\right)_{\partial K}=\left(z_{h,p}|^{intK},A_K\left(\vec{e}_v^{\,+}-\vec{e}_v|^{intK}\right)\cdot\vec{n}_K\right)_{\partial K}.$$

Let $A_K^{ml} = \delta_{ij}$ and $z_{h,p} = q_h^{ij}$. We get

$$\left\| q_h^{ij} \right\|^2 \le C \left\| q_h^{ij} \right\|_{\Gamma_h} \left\| e_v^j \right\|_{\Gamma_h} \le Ch^{-1} \left(\left\| \xi_v^j \right\| + h^{k+1} \right) \left\| q_h^{ij} \right\|.$$

It follows that

$$\left\| e_{\mathbb{P}}^{ij} \right\| = \left\| \frac{\partial e_v^j}{\partial x^i} + q_h^{ij} \right\| \le \left\| \frac{\partial \xi_v^j}{\partial x^i} \right\| + \left\| \frac{\partial \eta_v^j}{\partial x^i} \right\| + \left\| q_h^{ij} \right\| \le Ch^{-1} \left(\left\| \xi_v^j \right\| + \left\| \xi_v^i \right\| + h^{k+1} \right).$$

Thus we have

$$\| e_{\mathbb{P}} \| \le Ch^{-1} \left(\left\| \vec{\xi}_v \right\| + h^{k+1} \right).$$

According to (6.9), for small enough h, we have

$$|\mathcal{T}_1| \le \frac{\delta}{5} \left\| \vec{\xi}_v \right\|^2 + C \left\| \xi_u \right\|^2 + Ch^{2k+2}.$$

• **The estimate of $\mathcal{T}_2$.** For any given functions $f, \vec{g}$ in V_h^k, it follows

$$\begin{aligned}
&\sum_{K \in \mathcal{T}_h} \left\{ H_K^+ \left(\Lambda_K \vec{g}, f \right) + H_K^- \left(f, \Lambda_K \vec{g} \right) \right\} \\
&= -\sum_{K \in \mathcal{T}_h} \sum_{e_K \in \partial K} \left(f|^{intK}, \Lambda_K \vec{g}|^{intK} \cdot \vec{n}_{e_K} \right)_{e_K} + \sum_{K \in \mathcal{T}_h} \sum_{e_K \in \partial K} \left\{ \left(\Lambda_K \vec{g}^{\,+} \cdot \vec{n}_{e_K}, f|^{intK} \right)_{e_K} + \left(f^-, \Lambda_K \vec{g}|^{intK} \cdot \vec{n}_{e_K} \right)_{e_K} \right\}.
\end{aligned}$$

For each fixed edge e, we can find two neighbouring elements R and L such that $e = e_R = e_L = R \cap L$. Without loss of generality, we assume that $\vec{n}_{e_L} \cdot \vec{\beta} > 0$, then R is the "plus" side according to the definition in Section 2.2. Thus we have

$$\begin{aligned}
&- \left(f|^{intR}, \Lambda_R \vec{g}|^{intR} \cdot \vec{n}_{e_R} \right)_{e_R} + \left(\Lambda_R \vec{g}^{\,+} \cdot \vec{n}_{e_R}, f|^{intR} \right)_{e_R} + \left(f^-, \Lambda_R \vec{g}|^{intR} \cdot \vec{n}_{e_R} \right)_{e_R} \\
&= - \left(f^+, \Lambda_R \vec{g}^{\,+} \cdot \vec{n}_{e_R} \right)_e + \left(\Lambda_R \vec{g}^{\,+} \cdot \vec{n}_{e_R}, f^+ \right)_e + \left(f^-, \Lambda_R \vec{g}^{\,+} \cdot \vec{n}_{e_R} \right)_e = \left(\Lambda_R \vec{g}^{\,+} \cdot \vec{n}_{e_R}, f^- \right)_e,
\end{aligned}$$

and

$$\begin{aligned}
&- \left(f|^{intL}, \Lambda_L \vec{g}|^{intL} \cdot \vec{n}_{e_L} \right)_{e_L} + \left(\Lambda_L \vec{g}^{\,+} \cdot \vec{n}_{e_L}, f|^{intL} \right)_{e_L} + \left(f^-, \Lambda_L \vec{g}|^{intL} \cdot \vec{n}_{e_L} \right)_{e_L} \\
&= - \left(f^-, \Lambda_L \vec{g}|^- \cdot \vec{n}_{e_L} \right)_e + \left(\Lambda_L \vec{g}^{\,+} \cdot \vec{n}_{e_L}, f^- \right)_e + \left(f^-, \Lambda_L \vec{g}|^- \cdot \vec{n}_{e_L} \right)_e = \left(\Lambda_L \vec{g}^{\,+} \cdot \vec{n}_{e_L}, f^- \right)_e = - \left(\Lambda_L \vec{g}^{\,+} \cdot \vec{n}_{e_R}, f^- \right)_e.
\end{aligned}$$

Since F and u are smooth, we have

$$\sup_{\substack{e_R = e_L \\ R, L \in \mathcal{T}_h}} \left\{ \sup_{m,l=1,\dots,d} \left| \Lambda_R^{ml} - \Lambda_L^{ml} \right| \right\} \le Ch.$$

It follows that

$$\left| \sum_{K \in \mathcal{T}_h} \left\{ H_K^+ \left(\Lambda_K \vec{g}, f \right) + H_K^- \left(f, \Lambda_K \vec{g} \right) \right\} \right| \le Ch \left\| \vec{g} \right\|_{\Gamma_h} \| f \|_{\Gamma_h} \le C \left\| \vec{g} \right\| \| f \|.$$

Since ξ_u and $\vec{\xi}_v$ are in V_h^k, we have

$$|\mathcal{T}_2| \le C \left\| \vec{\xi}_v \right\| \| \xi_u \| \le \frac{\delta}{5} \left\| \vec{\xi}_v \right\|^2 + C \| \xi_u \|^2.$$

• **The estimate of $\mathcal{T}_3$.** We estimate $\mathcal{T}_3$ in two different cases.

Case 1: For the two-dimensional triangular meshes with P^k elements, we choose

$$\Pi^u = \Pi^v = \Pi.$$

According to the inverse inequality (6.3), we have

$$\begin{aligned}
|\mathcal{T}_3| &= \left| \sum_{K \in \mathcal{T}_h} \left\{ \left(\Lambda_K \vec{\eta}_v^{\,+} \cdot \vec{n}_K, \xi_u|^{intK} \right)_{\partial K} + \left(\eta_u^-, \Lambda_K \vec{\xi}_v|^{intK} \cdot \vec{n}_K \right)_{\partial K} \right\} \right| \\
&\le C \| \vec{\eta}_v \|_{\Gamma_h} \| \xi_u \|_{\Gamma_h} + C \| \eta_u \|_{\Gamma_h} \left\| \vec{\xi}_v \right\|_{\Gamma_h} \le \frac{\delta}{5} \left\| \vec{\xi}_v \right\|^2 + C \| \xi_u \|^2 + Ch^{2k}.
\end{aligned}$$

Case 2: For the two-dimensional Cartesian meshes with Q^k elements, we choose

$$\Pi^u = \Pi^-, \qquad \Pi^v = \Pi^+.$$

According to the property (6.1) of the projections $\Pi^{\pm}$, we have

$$H_K^+\left(A_K\vec{\eta}_v, f\right) = 0 \quad \text{and} \quad H_K^-\left(\eta_u, \vec{g}\right) = 0,$$

for any $f, \vec{g}$ in V_h^k and $A_K \in \mathbb{R}^{d\times d}$. Since ξ_u and $\Lambda_K\vec{\xi}_v$ are in V_h^k, we get $\mathcal{T}_3 \equiv 0$.

• **The estimate of $\mathcal{T}_4$.** Since u and F are smooth enough with bounded derivatives, we have

$$|\mathcal{T}_4| \le \frac{\delta}{5}\left\|\vec{\xi}_v\right\|^2 + C\left\|\xi_u\right\|^2 + Ch^{2k+2}.$$

Concluding above, we get

$$\int_D (\xi_u)_t\, \xi_u dx + \sum_{K\in\mathcal{T}_h}\int_K \vec{\xi}_v^{\,\mathrm{T}} \Lambda_K \vec{\xi}_v dx \le \frac{4\delta}{5}\left\|\vec{\xi}_v\right\|^2 + C\left\|\xi_u\right\|^2 + Ch^{2\rho},$$

where

$$\rho := \begin{cases} k, & \text{Case 1: triangular meshes with } P^k \text{ elements;} \\ k+1, & \text{Case 2: Cartesian meshes with } Q^k \text{ elements.} \end{cases}$$

Note that F is uniformly elliptic, i.e.,

$$F_{\mathbb{P}}\left(\Theta\right) := \left[\frac{\partial F}{\partial p^{ml}}\left(\Theta\right)\right]_{d\times d} \ge \delta I_{d\times d}.$$

We have $\Lambda_K \ge \delta I_{d\times d}$ for each $K \in \mathcal{T}_h$. It follows that

$$\frac{1}{2}\frac{d}{dt}\left\|\xi_u\right\|^2 + \frac{\delta}{5}\left\|\vec{\xi}_v\right\|^2 \le C\left\|\xi_u\right\|^2 + Ch^{2\rho}.$$

Then the Gronwall's inequality tells us that

$$\sup_{0\le t\le T}\left\|\xi_u(\cdot,t)\right\|^2 + \int_0^T \left\|\vec{\xi}_v(\cdot,t)\right\|^2 dt \le Ch^{2\rho}.$$

According to the triangle inequality, we finally get the error estimate for the LDG scheme

$$\sup_{0\le t\le T}\left\|u(\cdot,t) - u_h(\cdot,t)\right\| + \left(\int_0^T \left\|\nabla u(\cdot,t) - \vec{v}_h(\cdot,t)\right\|^2 dt\right)^{\frac{1}{2}} \le Ch^{\rho}. \tag{6.10}$$

Now, to complete the proof, let us verify the a priori assumption (6.8). In fact, the assumption in this theorem implies that $\rho \ge 4$. Thus the inequality (6.10) implies that the a priori assumption (6.8) is true. □

7. Conclusion

In this paper, we have developed an LDG method for directly solving fully nonlinear second-order PDEs (1.1) in one and multi-dimensional cases, in which the alternating numerical fluxes are used. With the help of local Gauss–Radau projection, we can prove the optimal error estimates ($\mathcal{O}(h^{k+1})$) for the Cartesian meshes with Q^k elements. By using the regular L^2-projection, we obtain the sub-optimal error estimates ($\mathcal{O}(h^k)$) for the triangular meshes with P^k elements. In particular, the optimal convergence rates can still be observed in a series numerical experiments when the P^k elements are used for both Cartesian and triangular meshes. Applications of our numerical schemes to stochastic optimal control problems are also discussed. We believe that our LDG method is promising in many other models in mathematical finance. As is well known, the viscosity solution to the fully nonlinear second-order PDEs is generically only Lipschitz continuous, with possibly discontinuous derivatives. Error estimates of the LDG method for such cases are much more difficult to analyse and are worthy of future investigation.

Data availability

Data will be made available on request.

Acknowledgements

The author would like to thank the editor, the associate editor, and two referees for their valuable comments and suggestions which led to a much improved version of the paper.

References

[1] L. Ambrosio, L.A. Caffarelli, Y. Brenier, G. Buttazzo, C. Villani, Optimal transportation and applications, Florence, Lecture Notes in Mathematics, vol. 1813, Springer-Verlag, Berlin; Centro Internazionale Matematico Estivo (C.I.M.E.), 2003, p. viii+164.
[2] X. Feng, R. Glowinski, M. Neilan, Recent developments in numerical methods for fully nonlinear second order partial differential equations, SIAM Rev. 55 (2) (2013) 205–267.
[3] M. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevskii, A reconstruction of the initial conditions of the universe by optimal mass transportation, Nature 417 (2002) 260–262.
[4] J. Yong, X.Y. Zhou, Stochastic Controls, in: Applications of Mathematics (New York), vol. 43, Springer-Verlag, New York, 1999, p. xxii+438, Hamiltonian systems and HJB equations.
[5] W. Reed, T. Hill, Triangular Mesh Methods for the Neutron Transport Equation, Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973, p. xix+530.
[6] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM J. Numer. Anal. 35 (6) (1998) 2440–2463.
[7] F. Bassi, S. Rebay, A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations, J. Comput. Phys. 131 (2) (1997) 267–279.
[8] B. Cockburn, S. Hou, C.-W. Shu, The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV: The multidimensional case, Math. Comp. 54 (190) (1990) 545–581.
[9] B. Cockburn, S.Y. Lin, C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III: One-dimensional systems, J. Comput. Phys. 84 (1) (1989) 90–113.
[10] B. Cockburn, C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II: General framework, Math. Comp. 52 (186) (1989) 411–435.
[11] B. Cockburn, C.-W. Shu, The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems, J. Comput. Phys. 141 (2) (1998) 199–224.
[12] M.G. Crandall, P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc. 277 (1) (1983) 1–42.
[13] C. Hu, C.-W. Shu, A discontinuous Galerkin finite element method for Hamilton-Jacobi equations, SIAM J. Sci. Comput. 21 (2) (1999) 666–690.
[14] F. Li, C.-W. Shu, Reinterpretation and simplified implementation of a discontinuous Galerkin method for Hamilton-Jacobi equations, Appl. Math. Lett. 18 (11) (2005) 1204–1209.
[15] Y. Cheng, C.-W. Shu, A discontinuous Galerkin finite element method for directly solving the Hamilton-Jacobi equations, J. Comput. Phys. 223 (1) (2007) 398–415.
[16] J. Yan, S. Osher, A local discontinuous Galerkin method for directly solving Hamilton-Jacobi equations, J. Comput. Phys. 230 (1) (2011) 232–244.
[17] T. Xiong, C.-W. Shu, M. Zhang, A priori error estimates for semi-discrete discontinuous Galerkin methods solving nonlinear Hamilton-Jacobi equations with smooth solutions, Int. J. Numer. Anal. Model. 10 (1) (2013) 154–177.
[18] X. Feng, T. Lewis, Mixed interior penalty discontinuous Galerkin methods for fully nonlinear second order elliptic and parabolic equations in high dimensions, Numer. Methods Partial Differential Equations 30 (5) (2014) 1538–1557.
[19] X. Feng, T. Lewis, Local discontinuous Galerkin methods for one-dimensional second order fully nonlinear elliptic and parabolic equations, J. Sci. Comput. 59 (1) (2014) 129–157.
[20] X. Feng, T. Lewis, Nonstandard local discontinuous Galerkin methods for fully nonlinear second order elliptic and parabolic equations in high dimensions, J. Sci. Comput. 77 (3) (2018) 1534–1565.
[21] S.C. Brenner, T. Gudi, M. Neilan, L.-y. Sung, $\mathscr{C}^0$ penalty methods for the fully nonlinear Monge-Ampère equation, Math. Comp. 80 (276) (2011) 1979–1995.
[22] S.C. Brenner, M. Neilan, Finite element approximations of the three dimensional Monge-Ampère equation, ESAIM Math. Model. Numer. Anal. 46 (5) (2012) 979–1001.
[23] Y. Cheng, Optimal error estimate of the local discontinuous Galerkin methods based on the generalized alternating numerical fluxes for nonlinear convection-diffusion equations, Numer. Algorithms 80 (4) (2019) 1329–1359.
[24] Y. Xu, C.-W. Shu, Error estimates of the semi-discrete local discontinuous Galerkin method for nonlinear convection-diffusion and KdV equations, Comput. Methods Appl. Mech. Engrg. 196 (37–40) (2007) 3805–3822.
[25] Q. Zhang, C.-W. Shu, Error estimates to smooth solutions of Runge-Kutta discontinuous Galerkin methods for scalar conservation laws, SIAM J. Numer. Anal. 42 (2) (2004) 641–666.
[26] J. Yan, C.-W. Shu, A local discontinuous Galerkin method for KdV type equations, SIAM J. Numer. Anal. 40 (2) (2002) 769–791.
[27] W. Zhao, T. Zhou, T. Kong, High order numerical schemes for second-order FBSDEs with applications to stochastic optimal control, Commun. Comput. Phys. 21 (3) (2017) 808–834.
[28] B. Cockburn, G. Kanschat, I. Perugia, D. Schötzau, Superconvergence of the local discontinuous Galerkin method for elliptic problems on cartesian grids, SIAM J. Numer. Anal. 39 (1) (2001) 264–285.
[29] P. Lasaint, P.-A. Raviart, On a finite element method for solving the neutron transport equation, in: Mathematical Aspects of Finite Elements in Partial Differential Equations (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1974), 1974, pp. 89–123, Publication No. 33.
[30] P. Castillo, B. Cockburn, D. Schötzau, C. Schwab, Optimal a priori error estimates for the *hp*-version of the local discontinuous Galerkin method for convection-diffusion problems, Math. Comp. 71 (238) (2002) 455–478.
[31] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, in: Studies in Mathematics and its Applications, Vol. 4, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978, p. xix+530.